

多参数 MRI 影像组学特征识别 HER-2 过表达型乳腺癌

薛珂, 李卓琳, 李振辉, 赵美琳, 孙诗昀, 丁莹莹

【摘要】 目的:探讨基于 DWI 和动态增强 MRI(DCE-MRI)的影像组学特征识别 HER-2 过表达型乳腺癌的可行性。**方法:**回顾性收集经穿刺或手术病理结果证实为乳腺浸润性导管癌的 380 例患者,在穿刺或手术前行 DWI 和 DCE-MRI 检查,手工逐层勾画 ADC 及 DCE 剪影图像上的病灶并提取感兴趣区的影像组学特征,利用秩和检验、Pearson 相关分析和支持向量机递归特征消除(SVM-RFE)进行特征筛选并降维,最后使用支持向量机(SVM)构建分类预测模型,采用受试者工作特征(ROC)曲线评估模型的预测效能。**结果:**HER-2 过表达型 75 例,非 HER-2 过表达型 305 例。从 ADC 及 DCE 剪影图像提取 1294 个特征,经过筛选及降维最终选出 6 个关键影像组学标签。SVM 构建的预测模型在训练集及验证集中的曲线下面积(AUC)分别为 0.79 及 0.81,均具有良好的诊断效能和一致性。**结论:**基于 DWI 和 DCE-MRI 影像组学特征构建的预测模型具有良好的稳定性及诊断效能,是识别 HER-2 过表达型乳腺癌的一种潜在手段。

【关键词】 乳腺肿瘤; 影像组学; 支持向量机(SVM); 磁共振成像; 扩散加权成像

【中图分类号】 R737.9; R445.2 **【文献标识码】** A **【文章编号】** 1000-0313(2020)02-0186-04

DOI:10.13609/j.cnki.1000-0313.2020.02.012

开放科学(资源服务)标识码(OSID):



Identify HER-2 over expression breast cancer based on radiomics of multi-parametric MRI XUE Ke, LI Zhuo-lin, LI Zhen-hui, et al. Department of Radiology, Yunnan Cancer Hospital, Kunming 650118, China

【Abstract】 Objective: The purpose of this study was to develop a predictive model based on diffusion weighted imaging (DWI) and dynamic contrast-enhanced MRI (DCE-MRI) radiomics to identify HER-2 over expression in breast cancer. **Methods:** 380 patients with invasive ductal carcinoma of breast confirmed by pathology were retrospectively analyzed. All underwent DWI and DCE-MRI scans before biopsy or surgery. The lesions on the ADC and DCE subtraction images were manually delineated and the radiomic histological characteristics of the region of interest were extracted. Feature selection and dimensionality reduction were conducted by use of Rank sum test, Pearson correlation analysis and SVM-RFE. SVM was eventually used to create the classification prediction model. Receiver operating characteristic curves (ROC) and areas under ROC were used to evaluate the diagnostic efficacy of the model. **Results:** There were 75 cases with HER-2 over expression and 305 cases with non-HER-2 over expression. 1294 features were obtained from ADC and DCE subtraction images and 6 radiomic features were finally selected after dimensionality reduction. The AUC of SVM model was 0.79 and 0.81 in training set and validation set, respectively, with good consistency and diagnostic efficiency. **Conclusion:** The predictive model based on ADC and DCE-MRI radiomics can identify HER-2 over expression in breast cancer. The multi-parametric MRI radiomic model is quite stable in terms of diagnostic efficiency.

【Key words】 Breast neoplasms; Radiomics; Support vector machine; Magnetic resonance imaging; Diffusion weighted imaging

作者单位: 650118 昆明, 云南省肿瘤医院(昆明医科大学第三附属医院)放射科

作者简介: 薛珂(1993—), 女, 云南建水人, 硕士研究生, 主要从事乳腺影像学诊断研究工作。

通讯作者: 丁莹莹, E-mail: dingyingying0428@163.com

基金项目: 云南省科技厅科技计划项目(省基础研究计划-昆医联合专项)(2018FE001-066); 云南省教育厅科学研究基金(研究生类)(2018Y056)

乳腺癌具有高度异质性, 不同分子亚型的乳腺癌其生物学行为、进展、治疗方案均不同。人类表皮生长因子受体 2(human epidermal growth factor receptor-2, HER-2)过表达型乳腺癌具有更高的侵袭性、复发

率和转移率,预后不佳。新辅助治疗的出现极大地改善这种状况,特别是以曲妥珠单抗为主的靶向治疗能提高患者的手术几率,降低复发及死亡风险。因此,早期识别 HER-2 过表达型乳腺癌有利于临床治疗方案的制定,以及延长患者无病生存及总生存时间。影像组学能从海量图像中挖掘肉眼无法识别的深层信息,提取并筛选 MRI 特征建立分类模型。本文利用影像组学提取并筛选 MRI-DWI 及动态增强(DCE)图像组学特征,并建立预测 HER-2 过表达型与非 HER-2 过表达型乳腺癌的分类模型,同时评估模型的预测效能。

材料与方法

1.临床资料

回顾性收集本院 2014 年 1 月—2016 年 12 月经穿刺或手术病理结果证实为乳腺浸润性导管癌患者共 400 例,并严格按照纳入及排除标准筛选入组病例。纳入标准:①在手术或穿刺前行乳腺 MRI 检查(所有患者在同一台设备上检查),包括 MRI 常规序列和动态增强扫描并剪影,扫描参数一致;②单侧单病灶乳腺癌;③肿块型病灶;④行免疫组化及 FISH 检测;⑤临床资料齐全,包括发病年龄、初潮年龄、绝经状态、怀孕次数、生产次数、淋巴结转移情况、肿瘤最大径等。排除标准:①既往有乳腺手术、化疗、放疗、激素或靶向治疗史患者;②图像质量不佳。最终 380 例患者纳入研究。

2.病理特征

根据 2013 年 St Gallen 国际乳腺癌会议专家共识^[1],通过免疫组化或荧光原位杂交得到雌激素受体(estrogen receptor, ER)、孕激素受体(progesterone receptor, PR)、细胞周期调节蛋白 Ki67 以及 HER-2 的表达情况,并将 HER-2 过表达型乳腺癌定义为 ER、PR 均表达阴性,而 HER-2 表达阳性。具有 3+ 的 HER-2 为阳性;当 HER-2 表达为 2+ 时(中等完全膜染色在 $\geq 10\%$ 的肿瘤细胞中),使用增强的原位杂交来检测 HER-2 扩增的存在(基因拷贝数 > 0.6 或 HER-2 与染色体 17 比率 > 2.2)。

3.MRI 图像采集

采用 Siemens Magnetom Avanto 1.5T 超导型 MRI 扫描仪和乳腺专用四通道相控阵表面线圈,患者俯卧位头先进,身体及双

肩放平,双侧乳腺自然悬垂于线圈内。DCE-MRI (TR 4.43 ms, TE 1.5 ms, 层厚 1.7 mm, 层间距 0.34 mm, 激励次数为 1) 共进行包括蒙片在内的 8 次重复扫描,每次扫描时间 60s。除蒙片与第 1 期增强扫描(即第 2 次扫描)间隔 24 s 为注药时间外,第 3~8 次扫描为连续无间隔。所获得的 9 个序列图像在 Siemens 工作站处理、分析。将增强图像与蒙片相减获得剪影图像。对比剂选用 Gd-DTPA(广州康臣药业, 15 mL/支),使用高压注射器经手背静脉团注,剂量为 0.2 mmol/kg,流率为 2.5 mL/s,蒙片扫描结束后立即注射对比剂,完毕后以相同流率注射 30 mL 生理盐水冲管。

4.影像组学分析与统计

图像获取:根据动态增强时间-信号强度曲线(TIC),在动态增强序列上找到病灶强化最明显时相的剪影图,将该剪影图与 ADC 图进行分析。

病灶分割:将 ADC 图及动态增强图像导入第三方软件(ITK-SNAP),由两位从事乳腺疾病诊断的影像医师在对应图像的每一个层面,沿着病灶边缘手动逐层勾画病灶,再经过融合得到三维感兴趣区(3D-ROI)(图 1)。

影像特征提取:使用 Matlab 软件提取 3D-ROI 的高通量特征,包括以下四类:①一阶统计特征:描述图像体素强度分布情况,如能量、熵、范围、偏度、均值、方差、标准差、均方根等;②形状大小特征:描述病灶三维空间信息,如最大径、表面积、体积、表面积与体积比、紧密度等;③纹理特征:描述灰度的空间相关性或体素强度的空间分布,提供图像上不同灰度级的相对位置信息,分别从灰度共生矩阵(GLCM)、灰度游程长度矩阵(GLRLM)特征进行纹理描述,如自突出、均匀

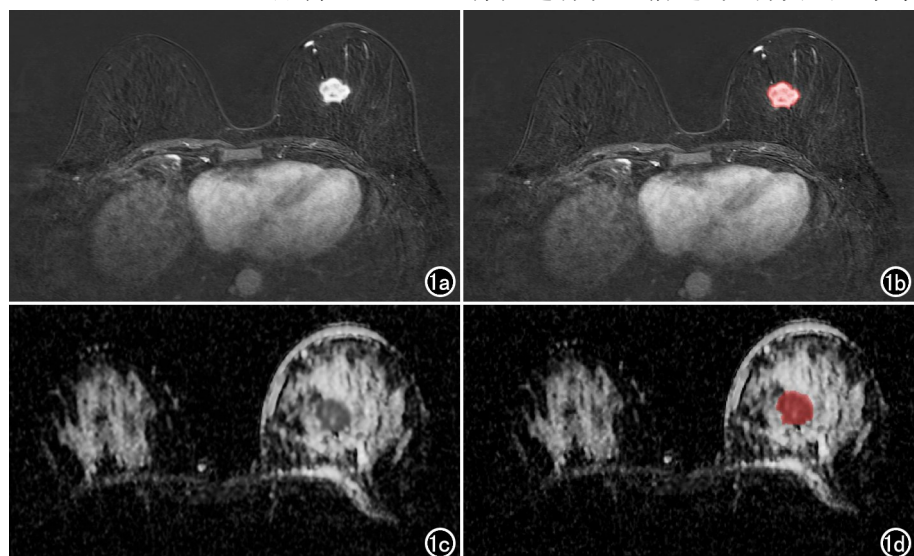


图 1 女,53 岁,左乳 HER-2 过表达型浸润性导管癌。a)DCE 剪影图显示病灶最大层面;b)DCE 剪影图上的 ROI 选择;c) ADC 图显示病灶最大层面;d)ADC 图上的 ROI 选择。

性、相关度、簇阴影等;④小波特征:小波变换能有效分解原始图像的纹理信息得到高频或低频的采样图像。

DCE 和 DWI 两个序列,每个序列提取 647 个影像特征(8 个形状大小特征、17 个一阶统计特征、54 个纹理特征和 568 个小波特征),共 1294 个特征。

特征筛选与预测模型构建:按患者的首次入院时间排序,前 2/3 作为训练集,后 1/3 作为验证集。首先采用秩和检验比较 HER-2 过表达型与非 HER-2 过表达型乳腺癌间的特征,筛选 $P<0.3$ 的特征进行下一步分析;然后将这些特征进行 Pearson 相关分析,计算相关系数(r),剔除 $r>0.85$ 的高度相关特征;最后利用支持向量机递归特征消除(support vector machine-recursive feature elimination,SVM-RFE)选择关键特征。RFE 将特征子集选择过程与分类器训练过程相结合,能充分衡量特征之间的相关性,有效削减特征的冗余,提升模型的预测效能和稳定性^[2],挑选出最小冗余、最大相关的最优特征集。

基于每个 ROI 的最优特征集,使用 SVM 方法构建最终的分类预测模型并评估其预测效果。考虑到有限的样本数,本研究采用十折交叉验证充分验证模型,并确定参数。使用受试者工作特征(ROC)曲线,计算曲线下面积(AUC)评估模型的预测效能。预测效能评价指标包括敏感性、特异性、准确性和 AUC。

结 果

共 380 例患者纳入研究,HER-2 过表达型 75 例,非 HER-2 过表达型 305 例。HER-2 过表达型中训练集 50 例,验证集 25 例;非 HER-2 过表达型中训练集 203 例,验证集 102 例。

最终从 1294 个影像特征中选出 6 个影像组学标签,分别为 Coif4_glszm_HGLZE、Coif2_glszm_LGLZE、N25_Coif6_glszm_HGLZE、Coif5_glrmlm_LRLGLE、Coif1_glrmlm_SRLGLE 和 ori_Max3D。

模型在训练集上的 AUC 为 0.79,准确率为 69.17%,敏感度为 81.97%,特异度为 65.1%,阳性预测值为 42.74%,阴性预测值 91.91%;在验证集上的 AUC 为 0.81,准确率为 67.71%,敏感度为 71.43%,特异度为 67.26%,阳性预测值为 21.27%,阴性预测值 95%(表 1、图 2)。

表 1 模型识别 HER-2 过表达型乳腺癌的效能

效能指标	训练集	验证集
AUC	0.79	0.81
准确率	69.17%	67.71%
敏感度	81.97%	71.43%
特异度	65.10%	67.26%

讨 论

HER-2 是一种原癌基因,是表皮生长因子受体(ERB)家族中的一员,这个家族的基因与细胞的生长、分化、转移过程调控作用有关,在许多肿瘤中均有异常表达^[3]。相对于 Luminal 型乳腺癌,HER-2 过表达型进展迅速,早期易发生淋巴结转移,复发、远处转移风险更高,预后更差^[4,5]。HER-2 是分子靶向药物的靶点及药效预测指标。分子靶向药物曲妥珠单抗可极大地改善 HER-2 阳性型乳腺癌的预后,被推荐为此类乳腺癌的标准辅助治疗方案^[6]。因此,及早识别 HER-2 阳性型乳腺癌有助于临床治疗方案的制定,以及降低患者的复发、死亡风险。

目前主要通过穿刺活检来获得肿瘤的分子分型,但是活检有创、且取材较局限,不能完全反映病灶的异质性。MRI 具有方便、无创、多功能成像的优点,能够反映病灶的形态及部分功能特征,在临床诊断与鉴别诊断、预测病理分子分型和评价疗效等方面具有巨大潜力。既往研究结果显示 DCE-MRI 能够反映病灶的形态学及血流动力学特征,并预测乳腺癌的分子亚型及远期预后^[7,8]。吴佩琪等^[9]的研究提示非肿块型乳腺癌的病灶 ADC 值与 HER-2 阳性表达呈正相关。DWI 在乳腺癌中的应用价值已获得广泛认可,肿瘤的 ADC 值越低,肿瘤细胞越密集,恶性可能性越大。乳腺癌的动态增强 TIC 一般呈快速流入型,表现为增强第 1 期明显强化,而剪影图像上剔除实质背景的强化,仅得到病灶的增强特征,故选取 ADC 图及剪影图像能反映更多的病灶信息。

然而,MRI 多依靠形态学特征进行诊断,缺乏一种客观的、定量的方法。影像组学能从影像图像中挖掘出高通量特征,识别肉眼无法直接获取的深层信息,并进行定量分析^[10-12]。马文娟等^[13]利用影像组学的方法提取乳腺 X 线摄影图像特征,所构建的预测模型

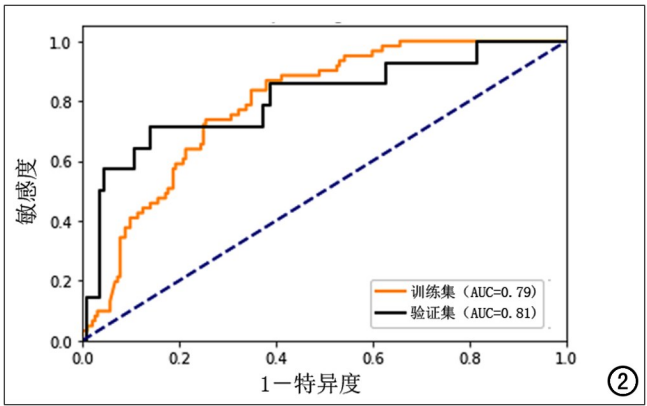


图 2 影像组学模型识别 HER-2 过表达型乳腺癌的 ROC 曲线。

的 AUC 为 0.806,能够很好地鉴别三阴性与非三阴性乳腺癌。Fan 等^[14]认为,联合 DCE-MRI 影像组学特征和临床信息能预测乳腺癌的分子分型,DCE-MRI 影像组学特征可作为鉴别乳腺癌四种分子亚型的候选标志物。杜鹏等^[15]基于多参数 MRI 影像组学特征,构建的膀胱癌复发的预测模型能在术前对膀胱癌的复发进行预测。

本研究提取 ADC 图及 DCE-MRI 剪影图像信息,再将所有特征汇总,利用秩和检验及 Pearson 相关分析筛选 HER-2 过表达组和非 HER-2 过表达组间的无意义特征及高度相关特征。SVM 是一种泛化能力良好、性能良好的二分类器,它先在低维空间计算出核函数,再将核函数映射到高维特征空间,得到最佳超平面^[16]。为了减少特征冗余,降低模型的过拟合程度,有效提高模型综合性能与泛化能力,进一步采用 SVM 选择最优特征子集^[16],最终得到 6 个影像组学标签,分别为 Coif4_glszm_HGLZE、Coif2_glszm_LGLZE、N25_Coif6_glszm_HGLZE、Coif5_glrml_LRLGLE、Coif1_glrml_SRLGLE 和 ori_Max3D。模型在训练集和验证集上的 AUC 值分别为 0.79 和 0.81,均具有很好的诊断效能;同时模型性能稳定,在训练集及验证集中 AUC 值接近,具有良好的一致性。本研究结果表明,基于 DCE-MRI 及 ADC 图像的多模态多参数 MRI 影像组学构建的预测模型能够有效地预测 HER-2 过表达型乳腺癌,有助于临床治疗方案的制定,为患者的个体化治疗提供依据。

本研究具有一定的创新性。多参数 MRI 可综合分析各序列特征,所建立的模型能充分结合序列的特点;SVM-RFE 消除冗余并建立预测模型,具有良好的诊断效能及泛化能力。本研究也存在一定局限性。病例数分到训练集及验证集数量偏少,且为单中心影像资料,需大量、多中心的病例资料验证模型的准确性。

综上所述,基于 DWI 和 DCE-MRI 影像组学特征能很好地区别 HER-2 过表达型与非 HER-2 过表达型乳腺癌,MRI 多参数影像组学模型具有良好的稳定性及诊断效能。

参考文献:

- [1] 王新昭,左文述,刘琪,等.2013 年 St Gallen 乳腺癌会议国际专家共识荟萃[J].中华肿瘤防治杂志,2013,20(23):1859-1864.

- [2] 南海燕,张欣,颜林枫,等.基于 DCE-MRI 定量参数的胶质瘤自动分级研究[J].磁共振成像,2018,9(7):494-499.
- [3] 郭晨明,吴楠,郭丽英.我国 HER-2 阳性乳腺癌预后影响因素的 Meta 分析[J].中国普通外科杂志,2016,25(2):264-270.
- [4] 赵玮,边莉,江泽飞.HER-2 基因突变在乳腺癌中的研究现状[J].中华乳腺病杂志(电子版),2016,10(6):362-365.
- [5] Parker JS,Mullins M,Cheang MC,et al.Supervised risk predictor of breast cancer based on intrinsic subtypes[J].J Clin Oncol,2009,27(8):1160-1167.
- [6] 江泽飞,邵志敏,徐兵河,等.人表皮生长因子受体 2 阳性乳腺癌临床诊疗专家共识 2016[J].中华医学杂志,2016,96(14):1091-1096.
- [7] 李逢芳,陈慧铀,王丽萍.Her-2 过表达型乳腺癌磁共振-扩散加权成像及动态增强成像分析[J].中国医疗设备,2016,31(10):51-54,61.
- [8] Pickles MD,Lowry M,Manton DJ,et al.Prognostic value of DCE-MRI in breast cancer patients undergoing neoadjuvant chemotherapy:a comparison with traditional indicators[J].Eur Radiol,2015,25(4):1097-1106.
- [9] 吴佩琪,赵可,吴磊,等.基于扩散加权成像和动态增强 MRI 的影像组学特征与乳腺癌分子分型的关系初探[J].中华放射学杂志,2018,52(5):338-343.
- [10] Li H,Zhu Y,Burnside ES,et al.Quantitative MRI radiomics in the prediction of molecular classifications of breast cancer subtypes in the TCGA/TCIA data set[J].NPJ Breast Cancer,2016,2:16012.
- [11] Mazurowski MA,Zhang J,Grimm LJ,et al.Radiogenomic analysis of breast cancer:Luminal B molecular subtype is associated with enhancement dynamics at MR imaging[J].Radiology,2014,273(2):365-372.
- [12] 艾国平,刘江勇,薛阳,等.DCE-MRI 定量参数对乳腺癌新辅助化疗的疗效评估及相关性分析[J].放射学实践,2018,33(11):44-49.
- [13] 马文娟,赵玉梅,季宇,等.基于乳腺 X 线影像组学特征的预测模型在鉴别三阴性与非三阴性乳腺癌中的价值[J].中华放射学杂志,2018,52(11):842-846.
- [14] Fan M,Li H,Wang S,et al.Radiomic analysis reveals DCE-MRI features for prediction of molecular subtypes of breast cancer[J].PLoS One,2017,12(2):e0171683.
- [15] 杜鹏,王焕军,石宇强,等.基于多参数磁共振成像影像组学特征的膀胱癌复发预测[J].中国医学装备,2018,15(11):46-50.
- [16] Anaissi A,Goyal M,Catchpoole DR,et al.Ensemble feature learning of genomic data using support vector machine[J].PLoS One,2016,11(6):e0157330.

(收稿日期:2019-05-29 修回日期:2019-09-04)